

Molecular Graphics-Structural and Molecular Graphics Descriptors in a QSAR Study of 17- α -Acetoxypregesterones

Rudolf Kiralj and Márcia M. C. Ferreira*

Instituto de Química, Universidade Estadual de Campinas, CP 6154, 13084-971 Campinas - SP, Brazil

Neste trabalho, foi feito um estudo de relações quantitativas entre a estrutura e a atividade biológica de 21 derivados de progesteronas ministrados via oral, dentre os quais 19 são 17- α -acetoxiprogesteronas. O método de quadrados mínimos parciais foi usado para construir modelos de regressão de boa qualidade, com $Q^2 = 0,707$ e $R^2 = 0,811$ utilizando duas componentes principais e quatro descritores. A maioria dos descritores moleculares foi obtida a partir de gráficos moleculares das geometrias otimizadas por meio de cálculos *ab initio* com um conjunto de base DFT 6-31G** (*descritores moleculares gráficos*). Os outros descritores foram obtidos pela combinação dos descritores anteriores com parâmetros estruturais experimentais extraídos do complexo progesterona-receptor da progesterona (*descritores gráfico-estruturais* ou *descritores gráficos e de modelagem*). Os modelos de regressão empregando somente cinco descritores gráficos e três componentes principais foram satisfatórios, $Q^2=0,556$, $R^2=0,718$, demonstrando a utilidade dos mesmos em estudos QSAR. Neste trabalho, onde foram estudados derivados de progesterona, ficou evidente que os descritores moleculares gráficos descreveram com sucesso os efeitos conformacionais, estéreo e eletrônicos dos diferentes substituintes.

Quantitative Structure-Activity Relationship study on 21 oral progestogens, 19 of which are 17- α -acetoxypregesterones, was performed by using Partial Least Squares. Fairly good regression models were achieved, the best being $Q^2=0.707$, $R^2=0.811$ with two Principal Components and four descriptors. Most of the molecular descriptors were generated from molecular graphics of DFT 6-31G** optimized geometries (*molecular graphics descriptors*) or were additionally combined with experimental structural parameters of progesterone receptor - progesterone complex (*molecular graphics-structural*, or *molecular graphics and modeling descriptors*). Regression models employing only these molecular graphics-based descriptors reached $Q^2=0.556$, $R^2=0.718$ with three Principal Components and five descriptors, demonstrating their usefulness in QSAR studies. In the case of progesterone derivatives, molecular graphics descriptors successfully included various conformational, steric and electronic substituent effects.

Keywords: progesterone derivatives, partial least squares, molecular graphics descriptors, molecular graphics-structural descriptors

Introduction

Progestogens, progesterone derivatives, nowadays are widely known as oral contraceptives. Besides that, health research¹⁻⁴ (hormone replacement and anti-cancer therapies, gynecological disorders etc.) and veterinary practice (animal birth control)^{5,6} are today the two most promising areas of progestogen applications. It is difficult to have an entirely clear picture of the progestogen behavior at atomic level due to the lack of large amount and homogeneity of progestogen activity data. Progesterone derivatives have

been used as target of various Structure-Activity Relationship (SAR) and Quantitative SAR (QSAR) studies since four decades ago.⁷ Progesterone (Figure 1), although having relatively simple structure, is a quite complicated molecular system. That is why researchers had to confront with the difficulty in quantifying the progestogen molecular properties without knowing 3D receptor-drug structure, and also to develop appropriate methods to treat the nonlinearity of steroid QSAR.⁸ Recently, the crystal structure of progesterone receptor (PR) – progesterone complex⁹ made possible to explain mutations at atomic level¹⁰ and perform more advanced drug design.

This work continues the SAR^{11,12} and QSAR¹³ idea to relate molecular descriptors of 21 oral progesterones to their

* e-mail: marcia@iqm.unicamp.br

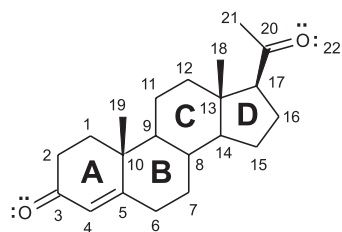


Figure 1. Progesterone nomenclature system.

oral progestational activity (relative to norethisterone, IC),¹⁴ 19 of which are 17- α -acetoxyprogesterones (Figure 2, Table 1), at the level of prediction using Partial Least Squares (PLS)¹⁵ regression models. In previous work¹³ various classes of molecular descriptors, as *a priori*,¹⁶ computed (at DFT *ab initio* level) and some molecular graphics-based descriptors were calculated, and PLS models were constructed and validated. It was observed that the new class of descriptors, molecular graphics-based descriptors, had significant contribution in PLS models, but more extensive study about this finding has not been performed yet. High-quality 2D projections of molecules or molecular aggregates obtained by current molecular graphics techniques can be an extensive source of quantitative information on molecular properties. In general, quantities directly “measured” from pictures using some digital or analogue technique can be 1D (linear, like molecular dimensions), 2D (surface areas of various molecular fragments projected onto the plane of projection or screen) or 3D (as molecular volume in some cases). Such measured descriptors, their combinations or functions can be named *molecular graphic descriptors*. Combination of these descriptors with some structural information from other sources (like data from experimental structure determination or molecular modeling), yields composite functions which can be called *molecular graphics-structural descriptors* or *molecular graphics and modeling descriptors*. Both classes of descriptors can be global (describing the entire molecule) or local (being related to some molecular fragment). Using

molecular graphics descriptors and some structural information from PR - progesterone complex modeling¹³ based on crystal structure of PR – progesterone complex,⁹ three sets of molecular graphics-structural descriptors were generated in this work. PLS models were built and validated for each data set, and the prediction of activity for three progestogens was performed (Figure 2). Finally, two composite descriptors, unweighted 3D-Morse signals 4 and 11,¹⁷ were added to the data set and new PLS models were constructed, validated and the predictions were performed. The meaning of the molecular-graphics based descriptors as well as their usefulness is discussed. The main goal of the regression analysis is to estimate the predictive power of PLS models based exclusively on molecular graphics-based descriptors, or on these and other types of molecular descriptors.

Methods

Molecular graphics descriptors

In previous work¹³ molecular geometries of progestogens **1-24** were optimized at DFT 6-31G** level, and high-quality figures of the molecules were constructed by positioning them along the C6(sp²)-substituent or C6(sp³)- α -substituent bond. Two projected surface areas were measured by analogue, empirical method¹⁶ as shown in Figure 3: the projected surface areas of substituent at C6 (S_6 – including, and S_6' - excluding hydrogens), and the projected surface area S of atoms or groups describing structural variations of the set **1-21** (H6 β and substituents at C1, C2, C6- β , C9-C13, C21). S_6' was set to zero for C6- α H atom. The choice of surfaces S_6 and S as molecular descriptors seemed to be reasonable as their structural variations are in accordance with the induced fit model.⁷

Most of substitutions in **1-21** are at position C6- α , and it was preliminary observed¹³ that the biological activity of this set of compounds is a quadratic function of the

Table 1. Progestogens under study (AOP=acetoxyprogesterone)

No.	Name	No.	Name
1	norethisterone (norethindrone)	13	6 α -fluoro-1-dehydro-17 α -AOP
2	17 α -AOP	14	6-fluoro-1,6-bisdehydro-17 α -AOP
3	17 α -ethinyltestosterone	15	6 α -methyl-1-dehydro-17 α -AOP
4	21-chloro-1,6-bisdehydro-17 α -AOP	16	6 α -chloro-1-dehydro-17 α -AOP
5	6 α -nitro-17 α -AOP	17	6-methyl-1,6-bisdehydro-17 α -AOP
6	6 β -chloro-17 α -AOP	18	6-methyl-6-dehydro-17 α -AOP
7	6 α -fluoro-17 α -AOP	19	6-fluoro-6-dehydro-17 α -AOP
8	21-fluoro-1,6-bisdehydro-17 α -AOP	20	6-chloro-1,6-bisdehydro-17 α -AOP
9	6 α -bromo-17 α -AOP	21	6-chloro-6-dehydro-17 α -AOP
10	6 α -methyl-17 α -AOP	22	18-nor-norethisteron
11	6 α -chloro-17 α -AOP	23	6-iodo-6-dehydro-17 α -AOP
12	6 α -bromo-1-dehydro-17 α -AOP	24	6-ethyl-6-dehydro-17 α -AOP

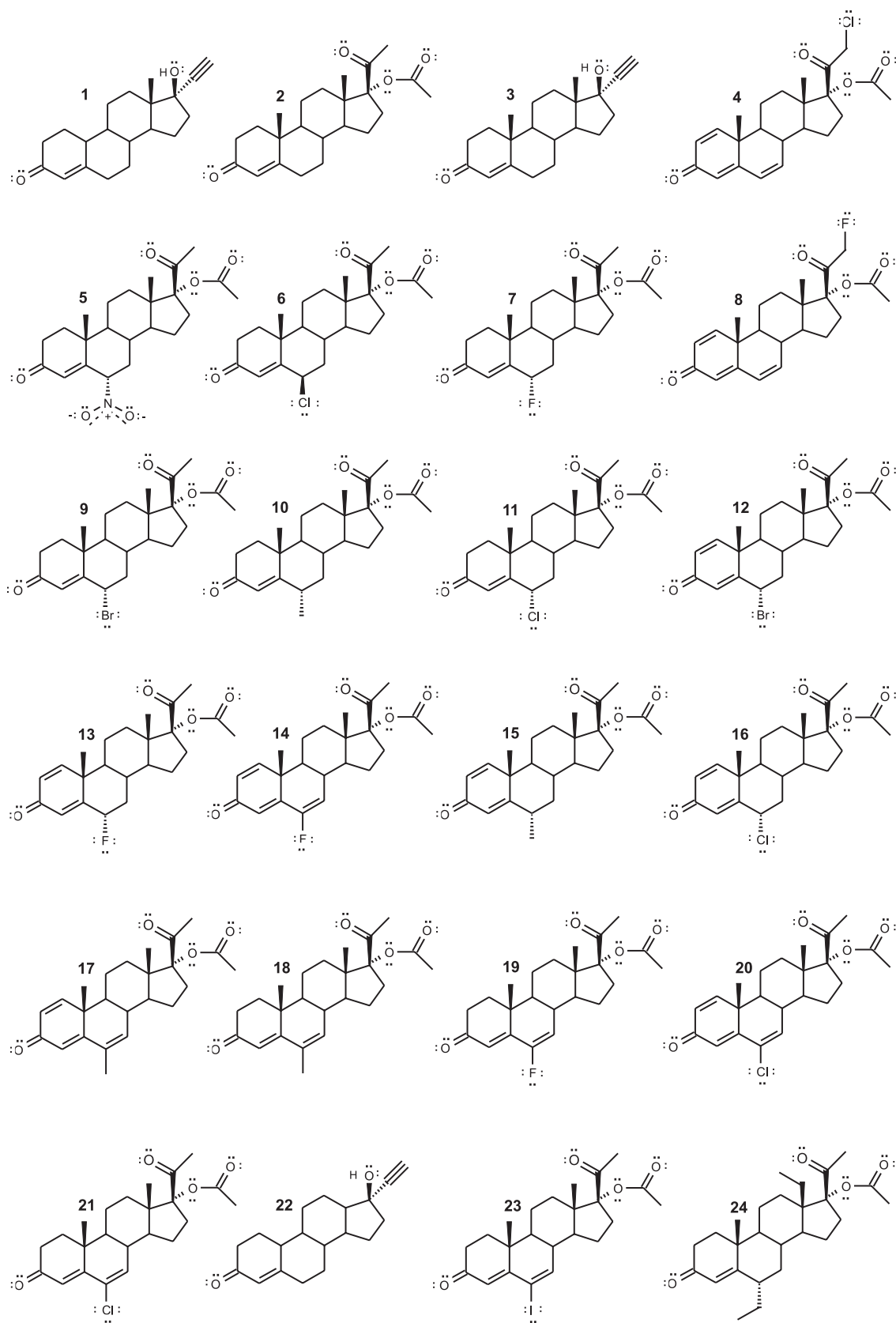


Figure 2. Schematic representation of progesterone derivatives under study, including multiple bonds and lone pairs. Training set: **1-21**; prediction set: **22-24**.

substituent size. On the other side, groups at positions C10, C13 or C21 also affect the activity. The third phenomenon to be taken into account is the saturation *e.g.* the presence of double bonds C1=C2, C4=C5 and/or C6=C7, as has been observed almost four decades ago.¹⁴ All these phenomena define the active conformation of the steroids, as Zeelen concluded more than two decades ago.¹⁸ Molecular graphics observations on studied compounds¹³ suggested that most of these structural changes can be better viewed along certain directions. The following composite molecular graphics descriptors were calculated:

$$P_1 = S_6 + S \quad (1)$$

$$P_2 = S_6' + S \quad (2)$$

$$P_3 = w_a S_6' + S \quad (3)$$

$$P_4 = w_b S_6 + S \quad (4)$$

where $w_a = c_3/c_2$, $w_b = d_3/d_2$, c_i and d_i are the coefficients from the linear regression equations $y = c_1 + c_2 S_6' + c_3 S$ and $y = d_1 + d_2 S_6 + d_3 S$, and y is the biological activity ($\log IC_{50}$). The descriptors S_6 , S , S_6' , $P_1 - P_4$ are presented in Table 2.

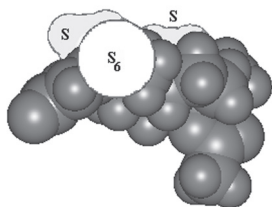


Figure 3. Definition of molecular graphics descriptors S_6 and S for molecule **23**.

Molecular graphics-structural descriptors

Preliminary molecular graphics and modeling study on PR - progesterone complexes¹³ including compounds **1-21** revealed that the nonlinear character of progesterone activity is mainly related to sterical relationships between the substituent at C6 (especially the substituent atom bound to C6) and sulfur atom from methionine 801 residue (Figure 4). The most appropriate substituents at C6 are Cl and CH_3 , while small (H, F) and big (Br) ones reduce the activity. Electronic relationships were not so clearly observed, but the polar NO_2 group which is of the appropriate size to fit in the hole between S(Met801) and C6 significantly reduces the activity with respect to that of **1**. S(Met801) participates in interactions with C7- H_2 group, which can be disturbed by large C6 substituents. On the other side, hydrophobic Met801 residue prefers to

interact with substituents of similar hydrophobicity, *i.e.* with non-polar or slightly polar groups such as CH_3 or higher halogens. The steric effects of C6 substituents were incorporated into new molecular graphics-structural descriptors in the following way: interatomic distances between S(Met801) and atoms of C6 substituent were measured;¹³ $D_{XS} - X...S$ distance (X is the substituent atom covalently bound to C6); $D_{YS} - Y...S$ distance (Y is the closest substituent atom to the Met801 sulfur); $D_{ZS} -$ equals to D_{XS} .

Van der Waals radii determined by Bondi¹⁸ were used: R_X - vdW radius of atom X, R_Y - vdW radius of Y, R_Z - equals to R_X with exception of CH_3 group (2.0 instead of 1.70 Å), and $R_S = 1.80$ Å for S. The measure of S-X,Y proximity was calculated as

$$\Delta_T = D_{TS} - (R_S + R_T) \quad (5)$$

where T=X, Y, Z. The three sets of D_{TS} , R_T values and experimental biological activities¹¹ are in Table 2. In this way, the substituent sterical effect was linearized. Six molecular graphics-structural descriptors were calculated for each data set:

$$M_{1T} = (w_a - |\Delta_T| / R_S) S_6' + S \quad (6)$$

$$M_{2T} = (w_a - \Delta_T / R_S) S_6' + S \quad (7)$$

$$M_{3T} = (w_b - |\Delta_T| / R_S) S_6 + S \quad (8)$$

$$M_{4T} = (w_b - \Delta_T / R_S) S_6 + S \quad (9)$$

$$P_{5T} = S (R_S - |\Delta_T|) / D_{TS} \quad (10)$$

$$P_{6T} = S (R_S - \Delta_T) / D_{TS} \quad (11)$$

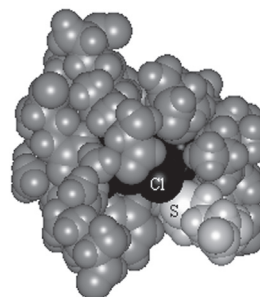


Figure 4. PR - 6α -chloro-progesterone complex at the active site of the protein. The closest amino-acid residues are shown. The drug (dark) was modeled by replacing 6α -H of progesterone by chlorine atom in the crystal structure of PR - progesterone complex (PDB code: 1A28⁹). Sulfur from methionine 801 (light) is in close contact with Cl atom of the drug.

Partial Least Squares models

Three data sets I, II and III were generated such that the descriptors S_6 , S , S_6' , $P_1 - P_4$ were common to all of them. In

Table 2. QSAR data for progestogens **1-24**

No.	$S_6/\text{\AA}^2$	$S/\text{\AA}^2$	$S_6'/\text{\AA}^2$	$D_{XS}/\text{\AA}$	$D_{YS}/\text{\AA}$	$D_{ZS}/\text{\AA}$	$R_X/\text{\AA}$	$R_Y/\text{\AA}$	$R_Z/\text{\AA}$	IC _{50,exp}
1	4.44	6.59	0.00	4.290	3.637	4.290	1.70	1.20	1.70	1.00
2	4.77	9.48	0.00	4.290	3.637	4.290	1.70	1.20	1.70	0.07
3	4.76	7.91	0.00	4.290	3.637	4.290	1.70	1.20	1.70	0.20
4	4.45	9.88	0.00	4.290	3.637	4.290	1.70	1.20	1.70	0.20
5	13.87	5.88	7.49	3.466	3.327	3.466	1.46	1.52	1.46	0.25
6	4.30	9.91	0.00	4.290	3.637	4.290	1.70	1.20	1.70	0.50
7	6.59	9.00	6.59	3.503	3.503	3.503	1.47	1.47	1.47	1.00
8	4.46	9.65	0.00	4.290	3.637	4.290	1.70	1.20	1.70	1.00
9	10.86	8.97	10.86	3.301	3.301	3.301	1.85	1.85	1.85	1.00
10	12.09	8.48	9.08	3.441	2.947	3.441	1.70	1.20	2.00	2.50
11	9.49	8.18	9.49	3.352	3.352	3.352	1.75	1.75	1.75	2.50
12	10.80	7.41	10.80	3.301	3.301	3.301	1.85	1.85	1.85	6.00
13	6.70	6.84	6.70	3.503	3.503	3.503	1.47	1.47	1.47	6.00
14	6.77	9.11	6.77	3.503	3.503	3.503	1.47	1.47	1.47	8.00
15	11.89	6.45	9.08	3.441	2.947	3.441	1.70	1.20	2.00	8.00
16	9.67	7.78	9.67	3.352	3.352	3.352	1.75	1.75	1.75	8.00
17	12.10	8.94	9.08	3.441	2.947	3.441	1.70	1.20	2.00	10.00
18	12.17	10.01	9.08	3.441	2.947	3.441	1.70	1.20	2.00	12.00
19	6.70	10.12	6.70	3.503	3.503	3.503	1.47	1.47	1.47	15.00
20	9.65	9.41	9.65	3.352	3.352	3.352	1.75	1.75	1.75	35.00
21	9.55	10.78	9.55	3.352	3.352	3.352	1.75	1.75	1.75	50.00
22	4.41	3.24	0.00	4.290	3.637	4.290	1.70	1.20	1.70	-
23	12.37	10.18	12.37	3.247	3.019	3.247	1.98	1.98	1.98	-
24	17.94	8.03	14.14	3.441	2.947	3.441	1.70	1.20	2.00	-

addition, the descriptors $M_{IT} - M_{4T}$, P_{5T} , P_{6T} (for T = X, Y, and Z) were calculated for data sets I, II and III, respectively. Variable selection, validation (leave-one out crossvalidation) of the models and prediction were performed using Pirouette software.²⁰ The data sets were treated independently. The main purpose of this data analysis was to evaluate the binding power and usability of molecular graphics and molecular graphics-structural descriptors for the progestogen QSAR. Finally, knowing that these descriptors cannot describe entirely the 17-acetoxypregesterone activity, two more descriptors previously calculated¹³ - unweighted 3D-Morse signals 4 (M_{04}) and 11 (M_{11}),¹⁷ were added to each data set. The final models were compared to molecular graphics-based models.

Results and Discussion

Correlations of all molecular descriptors with the biological activity (log IC) are presented in Table 3. PLS models for data sets I, II and III and predictions for **22-24** are in Table 4; the models are named (a) when using all molecular graphics-based descriptors, (b) after variable selection of these descriptors, (c) all descriptors from (a) plus two 3D-Morse signal descriptors, (d) models analogous to the best model from our previous work¹³ which is presented in Figure 5 (Id).

Table 3. Correlation coefficients between the biological activity (log IC) and molecular descriptors

S_6	0.448	M_{3X}	0.631	P_{6Y}	0.703
S	0.168	M_{3X}	0.640	M_{1Z}	0.535
S_6'	0.708	P_{5X}	0.786	M_{2Z}	0.643
P_1	0.536	P_{6X}	0.755	M_{3Z}	0.671
P_2	0.752	M_{1Y}	0.596	M_{4Z}	0.656
P_3	0.732	M_{2Y}	0.612	P_{5Z}	0.788
P_4	0.743	M_{3Y}	0.561	P_{6Z}	0.778
M_{1X}	0.608	M_{4Y}	0.594		
M_{2X}	0.621	P_{5Y}	0.671		

The correlation coefficients from Table 3 reveal interesting structure-activity relationships. Correlations concerning S , S_6 , S_6' are low, moderate and high, respectively. By other words, the activity for **1-21** is determined more by C6 substituents than by any other (at C18, C19, C21 substitution sites), what is expected since the active site hole of PR, even after its complexation with progesterone, has the largest unoccupied space around C6.¹³ Exclusion of hydrogens in S_6' showed to be even a better choice, as they could be considered soft atoms, and so methyl and ethyl groups can be approximated as one or two carbon atoms. Weighted and unweighted linear combinations of descriptors $P_1 - P_4$ (equations 1-4) result in better descriptors, what is in accordance with above mentioned induced fit model when all the changes (substitutions, saturations, etc.)

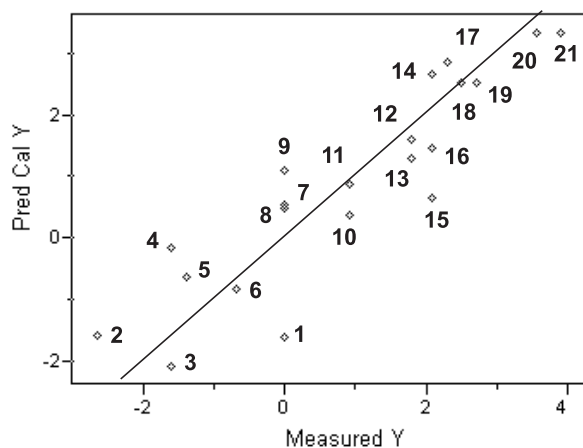
Table 4. PLS models and predictions for 22-24

Model ^a	Molecular Descriptors	PC	R ²	Q ²	SEP ^b	22	23	24
Ia	S ₆ ['] , S, S ₆ ['] , P ₁ - P ₆ , M ₁ - M ₄	3	0.685	0.469	1.326	-1.33	1.87	1.19
Ib	S ₆ ['] , P ₁ , P ₃ , P ₅ , P ₆	2	0.687	0.525	1.232	-1.46	2.19	0.29
Ic	S ₆ ['] , S, S ₆ ['] , P ₁ - P ₆ , M ₁ - M ₄ , M ₀₄ , M ₁₁	4	0.864	0.671	1.045	0.40	-1.20	3.18
Id ^c	S ₆ ['] , P ₃ , M ₀₄ , M ₁₁	2	0.811	0.707	0.966	-0.92	0.23	3.02
IIa	S ₆ ['] , S, S ₆ ['] , P ₁ - P ₆ , M ₁ - M ₄	3	0.707	0.483	1.307	-1.76	0.08	1.25
IIb	S ₆ ['] , P ₁ , P ₃ , P ₅ , P ₆	3	0.718	0.556	1.190	-1.10	1.59	0.10
IIc	S ₆ ['] , S, S ₆ ['] , P ₁ - P ₆ , M ₁ - M ₄ , M ₀₄ , M ₁₁	4	0.869	0.688	1.015	0.54	-1.76	3.34
IId	S ₆ ['] , P ₃ , M ₀₄ , M ₁₁	3	0.826	0.714	0.956	-1.49	0.00	4.22
IIIa	S ₆ ['] , S, S ₆ ['] , P ₁ - P ₆ , M ₁ - M ₄	3	0.652	0.389	1.424	-1.51	2.56	1.07
IIIb	S ₆ ['] , P ₁ , P ₃ , P ₅ , P ₆	2	0.660	0.513	1.248	-0.75	2.18	0.45
IIIc	S ₆ ['] , S, S ₆ ['] , P ₁ - P ₆ , M ₁ - M ₄ , M ₀₄ , M ₁₁	3	0.836	0.636	1.090	0.39	0.40	3.02
IIId	S ₆ ['] , P ₃ , M ₀₄ , M ₁₁	2	0.806	0.697	0.982	-0.37	0.35	2.70

^aIndexes X, Y, Z for variables M₁-M₄, P₅, P₆ were omitted as the variables were named for each data set I, II, III;

^bStandard error of prediction for validation;

^cThe best model in previous work.¹³

**Figure 5.** PLS plot for model Id.

should determine the active conformation of a progesterone. M₁ - M₄ descriptors (equations 6-9) represent extended weighting scheme for linear combination of S and S₆[']; variable D includes information about the proximity between sulfur from Met801 residue and some substituent atom. If the distance between the atoms is much greater than van der Waals sum of the atomic radii (usually with 0.2 Å tolerance), the atoms have no contact and so no interaction occur. If the atoms penetrate to each other beyond this tolerance, electron correlation would set them at an equilibrium distance, interfering into other drug-protein interactions; by other words, the steric effect would reduce the activity. This substituent size-activity relationship showed to be strictly quadratic at substitution sites C6 and C21.¹³ Thus the absolute value of Δ (equation 5) would linearize this effect. M₁ - M₄ seem not to be better than S₆['] and P₁ - P₄ (Table 3). Inclusion of Δ into P₅ and P₆ (equations

10 and 11) give new variables which are much better than P₁ - P₄ and even more suitable than S₆['] (correlation coefficients reaching 0.79). One could conclude *a priori* that S₆['], P₅ and P₆ would be the best variables for QSAR models, but more precise study on the subject should be carried out in variable selection.

The PLS study presented in Table 4 suggests which descriptors would be the most appropriate for progesterone QSAR, and also that models with good quality can be achieved using only these molecular graphics-based descriptors. Two cases should be distinguished among the twelve models: models based only on molecular graphics and molecular graphics-structural descriptors, and models including 3D-Morse variables. In the first case, data compression yields three significant Principal Components (models Ia, IIa, IIIa) and reaches R=0.841, Q=0.695 (model IIa). After variable selection further compression is achieved for Ib and IIIb, while not for IId; these models include two more variables than the models just suggested above (P₁ and P₃ show to be the best descriptors, Table 3). These models are not much better than the previous ones (R=0.847, Q=0.746), but nicely illustrate the meaning and usefulness of molecular graphics-based descriptors: five of them are necessary for a QSAR model, and they can describe the PR - progesterone binding as 2D phenomena. These facts encourage the search for other types of descriptors which can bring some new information on molecular properties, so new PLS models would be more quantitative. Models Ic, IIc, IIIc are such an attempt. The best R=0.932, Q=0.829 (IIc) and standard error of prediction (SEP) was reduced with respect to the previous models. On the other side, these models use three or more PCs and all the variables. Descriptors M₀₄ and M₁₁ contain some information in common with molecular graphics-based descriptors, so only

four descriptors (S_6' , P_5 , M_{04} , M_{11}) are sufficient to build a good PLS model (Id, IId, IIIId) with only two PCs (the best $R=0.909$, $Q=0.845$ in model IId). These three models seem to be the best ones in QSAR studies.

Prediction of activities for **22-24**, compared to the expected¹³ (**22** – non-active due to lack of Me at C18 and C19; **23** – less active than its chlorine analogue; **24** – highly active due to Et placed left to S from Met801), is an additional criterion for searching the best models. Molecule **22** should be even less active than **2** ($IC < 0.07$), the activity of **23** in between that of **19** and **21** ($IC = 23$ to 50), and **24** is expected to be far more active than **10** ($IC > 3$) and even more than **21** ($IC > 50$). Models Ic, IId and IIIc predicted the activities in decreasing order **24** – **22** – **23** instead of **24** – **23** – **22** as expected. Considering all the parameters, Id (Figure 5), IId and IIIId can be used as the best PLS models. Of course, in the case where only molecular graphics based-descriptors are utilized, Ia, IIa and IIIa are recommendable models.

Conclusions

New molecular graphics-based descriptors were defined and calculated for 24 progestogens in a QSAR study. Biological activities (oral progestational activities relative to norethisterone) were calculated for **1-21** and predicted for **22-24** employing various PLS models. The best PLS models include molecular graphics-based and 3D-Morse descriptors, and reproduce biological activities for **1-21** satisfactorily well. PLS models show that molecular graphics and molecular graphics-structural descriptors, although having prevalent contribution, are not sufficient to build a high-quality PLS model. The chemical meaning of molecular graphics-based descriptors for progestogens is fully understandable in terms of induced fit model. Prediction of activity for **22-24** using the best models is in accord with expectations.

Acknowledgment

The authors acknowledge FAPESP for the financial support.

References

1. Gambrell, R. D., Jr.; *Arch. Family Med.* **1996**, *5*, 341.
2. Gambrell, R. D., Jr.; *Int. J. Fertility* **1989**, *34*, 315.
3. Desport, J. C.; Blanc-Vincent, M. P.; Gory-Delabaere, G.; Bachmann, P.; Beal, J.; Benamouzig, R.; Colomb, V.; Kere, D.; Melchior, J. C.; Nitenberg, G.; Raynard, B.; Scheiner, S.; Senesse, P.; *Bull. Cancer* **2000**, *87*, 315.
4. Baziad, A.; Pache, T.; *First Consensus Meeting on Menopause in the East Asian Region*, Geneva, Switzerland, 1997; <http://matweb.hcuge.ch/matweb/bookmp/>, accessed: 6 August 2001.
5. Kustritz, M. V. R. In *Recent Advances in Small Animal Reproduction*; Concannon, P. W.; England, G.; Verstegen, J., eds.; International Veterinary Information Service: Ithaca; New York. www.ivis.org, accessed: 6 August 2001.
6. Löfstedt, R., In *VETGATE, The UK's Gateway to High Quality Internet Resources in Animal Health*. <http://www.upei.ca/~lofstedt/opence/pande.html>, accessed: 6 August 2001.
7. Bursi, R.; Groen, M. B.; *Eur. J. Med. Chem.* **2000**, *35*, 787.
8. So, S.-S.; van Helden, S. P.; Geerestein, J. V.; Karplus, M.; *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 762.
9. Williams, S. P.; Sigler, P. B.; *Nature* **1998**, *393*, 392.
10. Matias, P. M.; Donner, P.; Coelho, T.; Thomaz, M.; Peixoto, C.; Macedo, S.; Otto, N.; Joschko, S.; Scholz P.; Wegg, A.; Bäsler, S.; Schäfer, M.; Egner, U.; Carrondo, M. A.; *J. Biol. Chem.* **2000**, *275*, 26164.
11. Vendrame, R.; M. Ferreira, M. M. C.; Collins, C. H.; Takahata, Y.; *J. Mol. Graph. Mod.* **2002**, *20*, 345.
12. Braga, R. S.; Vendrame, R.; Galvão, D. S.; *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1377.
13. Kiralj, R.; Takahata, Y.; Ferreira, M. M. C.; *Quant. Struct.-Act. Relat.*, accepted for publication.
14. Shoppe, C. W.; *Chemistry of the steroids*, 2nd ed., Butterworth: London, 1964, p. 178.
15. Beebe, K. R.; Pell, R. J.; Seasholtz, M. B.; *Chemometrics: A Practical Guide*, Wiley: New York, 1998.
16. Kiralj, R.; Ferreira, M. M. C.; *J. Mol. Graph. Mod.*, in press.
17. Schuur, J.; Gasteiger, J.; *Anal. Chem.* **1997**, *69*, 2398.
18. Zeelen, B. J.; *Biol. Act. Chem. Struct.* **1977**, 147.
19. Bondi, A.; *J. Phys. Chem.* **1964**, *68*, 441.
20. Pirouette 3.01, Infometrix, Inc., Seattle, WA, 2001.

Received: November 11, 2001

Published on the web: November 19, 2002

FAPESP helped in meeting the publication costs of this article.